Fig.1

Constructs for secretion of *S. pombe* hydrophobins

Construct

Signal sequence

Authentic protein

P factor signal peptide

PDewA

P factor signal peptide +
6 amino acids

P+6DewA

P factor signal peptide
including P-factor

PfakDewA

↑ = postulated protease cleavage sites
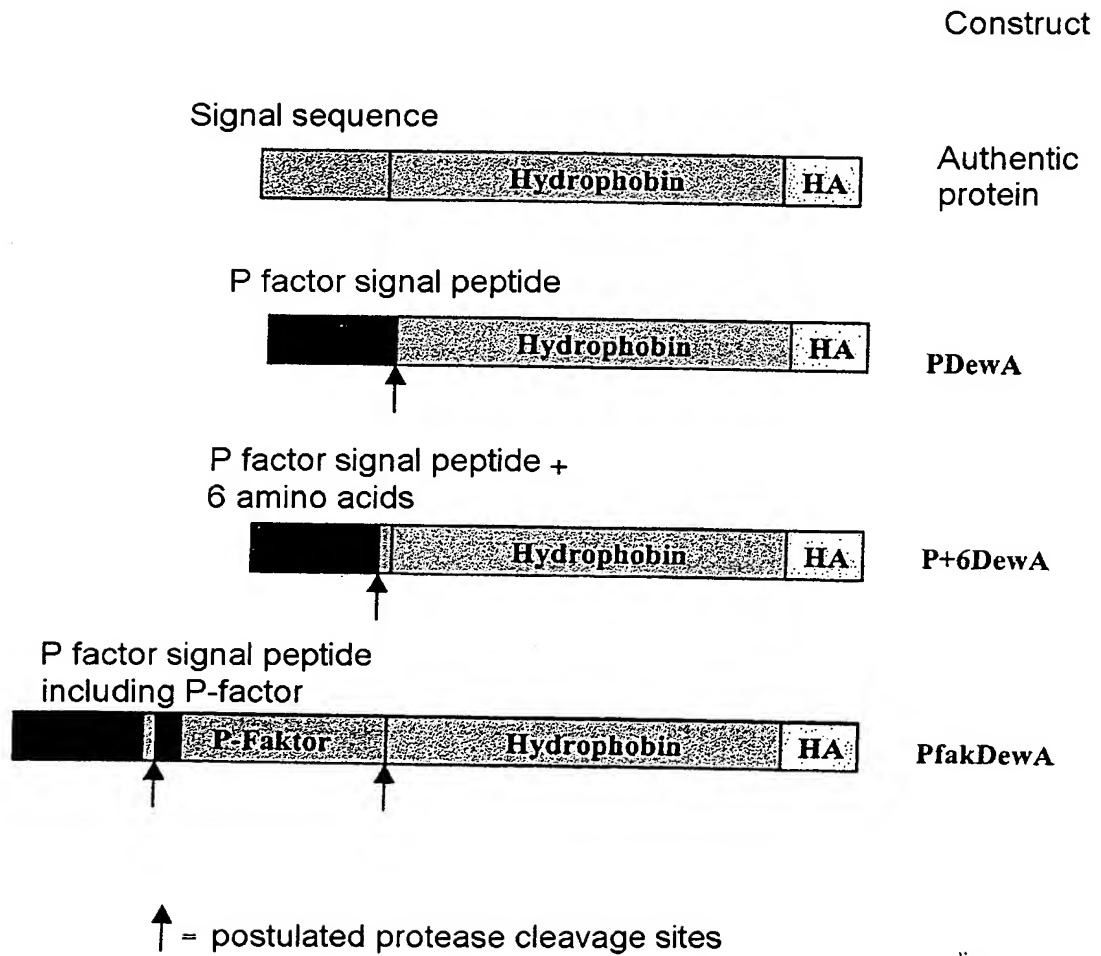
Fig. 2

A) Genomic sequence of the DewA gene:
(the sequenes of the two introns are underlined)

```
ATGCGCTTCA TCGTCTCTCT CCTCGCCTTC ACTGCCGCGG CCACCGCAAC CGCCCTCCCG
GCCTCTGCCG CAAAGAACGC GAAGCTGGCC ACCTCGGCGG CCTTCGCCAA GCAGGCTGAA
GGCACCACCT GCAATGTCGG CTCGATCGCT TGCTGCAACT CCCCCGCTGA GACCAACAAC
GACAGTCTGT TGAGCGGTCT GCTCGGTGCT GGCCTTCTCA ACGGGCTCTC GGGCAACACT
GGCAGCGCCT GCGCCAAGGC GAGCTTGATT GACCAGCTGG GTCTGCTCGG TACGTGATCC
CCACTCAGTC GCTCCCGGAG AGGCTGAGGG AAGACGAGCG ACGGTCTAGA AATGGTGTGC
TAATAGATGC ATGTGTGCAG CTCTCGTCGA CCACACTGAG GAAGGCCCCG TCTGCAAGAA
CATCGTCGCT TGCTGCCCTG AGGGAACCAC CAACGTACGT CTTTCAGATC TGCTACAAGT
GAGGCGATCA AAACTAACAT ATTCCAGTGT GTTGCCGTCG ACAACGCTGG CGCCGGTACC
AAGGCTGAGT AA
```

B) Sequence of the *Aspergillus nidulans* DewA protein:

```
MRFIVSLLAF TAAATATALP ASAAKNAKLA TSAAFAKQAE GTTCNVGSIA CCNSPAETNN
DSLLSGLLGA GLLNGLSGNT GSACAKASLI DQLGLLALVD HTEEGPVCKN IVACCPEGTT
NCVAVDNAGA GTKAE
```

```
(ATGCGCTTCA TCGTCTCTCT CCTCGCCTTC ACTGCCGCGG CCACCGCAAC CGCCCTCCCG
GCCTCTGCCG CAAAGAACGC GAAGCTGGCC ACCTCGGCGG CCTTCGCCAA GCAGGCTGAA
GGCACCACCT GCAATGTCGG CTCGATCGCT TGCTGCAACT CCCCCGCTGA GACCAACAAC
GACAGTCTGT TGAGCGGTCT GCTCGGTGCT GGCCTTCTCA ACGGGCTCTC GGGCAACACT
GGCAGCGCCT GCGCCAAGGC GAGCTTGATT GACCAGCTGG GTCTGCTCGC TCTCGTCGAC
CACACTGAGG AAGGCCCCGT CTGCAAGAAC ATCGTCGCTT GCTGCCCTGA GGGAACCACC
AACTGTGTTG CCGTCGACAA CGCTGGCGCC GGTACCAAGG CTGAGTAA)
```

C) Ha-Tag sequence:

```
LVPRGSIEGR GGRIFYPYDV PDYAGYPYDV PDYAGSYPYD VPDYAAQCGR
(CTGGT    TCCGCGTGGA TCCATCGAAG GTCGTGGCGG CCGCATCTTT   TACCCATACG
ATGTTCCTGA CTATGCGGGC TATCCCTATG ACGTCCCGGA CTATGCAGGA TCCTATCCAT
ATGACGTTCC AGATTACGCT GCTCAGTGCG GCCGCTAATA G)
```

Fig.3

A) Sequence of the P-factor pre-protein:

MKITAVIALL   FSLAAASPIP   VADPGVVSVS   KSYADFLRVY   QSWNTFANPD   RPNLKKREFE

AAPAKTYADF   LRAYQSWNTF   VNPDRPNLKK   REFEAAPEKS   YADFLRAYHS   WNTFVNPDRP

NLKKREFEAA   PAKTYADFLR   AYQSWNTFVN   PDRPNLKKRT   EEDEENEEED   EEYYRFLQFY

IMTVPENSTI   TDVNITAKFE   S

(ATGAAGATCA   CCGCTGTCAT   TGCCCTTTTA   TTCTCACTTG   CTGCTGCCTC   ACCTATTCCA

GTTGCCGATC   CTGGTGTGGT   TTCAGTTAGC   AAGTCATATG   CTGATTTCCT   TCGTGTTTAC

CAAAGTTGGA   ACACTTTTGC   TAATCCTGAT   AGACCCAACT   TGAAAAAGCG   CGAATTCGAA

GCTGCTCCCG   CAAAAACTTA   TGCTGATTTC   CTTCGTGCTT   ATCAAAGTTG   GAACACTTTT

GTTAATCCTG   ACAGACCCAA   TTTGAAAAAG   CGTGAGTTTG   AAGCTGCCCC   AGAGAAGAGT

TATGCTGATT   TCCTTCGTGC   TTACCATAGT   TGGAACACTT   TTGTTAATCC   TGACAGACCC

AACTTGAAAA   AGCGCGAATT   CGAAGCTGCT   CCCGCAAAAA   CTTATGCTGA   TTTCCTTCGT

GCTTACCAAA   GTTGGAACAC   TTTTGTTAAT   CCTGACAGAC   CCAACTTGAA   AAAGCGCACT

GAAGAAGATG   AAGAGAATGA   GGAAGAGGAT   GAAGAATACT   ATCGCTTTCT   TCAGTTTTAT

ATCATGACTG   TCCCAGAGAA   TTCCACTATT   ACAGATGTCA   ATATTACTGC   CAAATTTGAG

AGCTAA)

B) Sequence of the removable signal peptide and of the P-factor pre-protein 6 amino acids downstream thereof:

MKITAVIALL   FSLAAASPIP   VADPGV

(ATGAAGATCA   CCGCTGTCAT   TGCCCTTTTA   TTCTCACTTG   CTGCTGCCTC   ACCTATTCCA
GTTGCCGATC   CTGGTGTG)

C) Sequence utilized for „P shuttle":

MKITAVIALL   FSLAAASPIP   VADPGVVSVS   KSYADFLRVY   QSWNTFANPD   RPNLKKR

(ATGAAGATCA   CCGCTGTCAT   TGCCCTTTTA   TTCTCACTTG   CTGCTGCCTC   ACCTATTCCA
GTTGCCGATC   CTGGTGTGGT   TTCAGTTAGC   AAGTCATATG   CTGATTTCCT   TCGTGTTTAC
CAAAGTTGGA   ACACTTTTGC   TAATCCTGAT   AGACCCAACT   TGAAAAAGCG   C)

Fig. 4

Fusion protein comprising the „P-shuttle" sequence, the mature DewA and the C-terminally fused HA-Tag:

MKITAVIALL FSLAAASPIP VADPGVVSVS KSYADFLRVY QSWNTFANPD RPNLKKRLPA

SAAKNAKLAT SAAFAKQAEG TTCNVGSIAC CNSPAETNND SLLSGLLGAG LLNGLSGNTG

SACAKASLID QLGLLALVDH TEEGPVCKNI VACCPEGTTN CVAVDNAGAG TKAELVPRGS

IEGRGGRIFY PYDVPDYAGY PYDVPDYAGS YPYDVPDYAA QCGR

(ATGAAGATCA CCGCTGTCAT TGCCCTTTTA TTCTCACTTG CTGCTGCCTC ACCTATTCCA
GTTGCCGATC CTGGTGTGGT TTCAGTTAGC AAGTCATATG CTGATTTCCT TCGTGTTTAC
CAAAGTTGGA ACACTTTTGC TAATCCTGAT AGACCCAACT TGAAAAAGCG CCTCCCGGCC
TCTGCCGCAA AGAACGCGAA GCTGGCCACC TCGGCGGCCT TCGCCAAGCA GGCTGAAGGC
ACCACCTGCA ATGTCGGCTC GATCGCTTGC TGCAACTCCC CCGCTGAGAC CAACAACGAC
AGTCTGTTGA GCGGTCTGCT CGGTGCTGGC CTTCTCAACG GGCTCTCGGG CAACACTGGC
AGCGCCTGCG CCAAGGCGAG CTTGATTGAC CAGCTGGGTC TGCTCGCTCT CGTCGACCAC
ACTGAGGAAG GCCCCGTCTG CAAGAACATC GTCGCTTGCT GCCCTGAGGG AACCACCAAC
TGTGTTGCCG TCGACAACGC TGGCGCCGGT ACCAAGGCTG AGCTGGTTCC GCGTGGATCC
ATCGAAGGTC GTGGCGGCCG CATCTTTTAC CCATACGATG TTCCTGACTA TGCGGGCTAT
CCCTATGACG TCCCGGACTA TGCAGGATCC TATCCATATG ACGTTCCAGA TTACGCTGCT
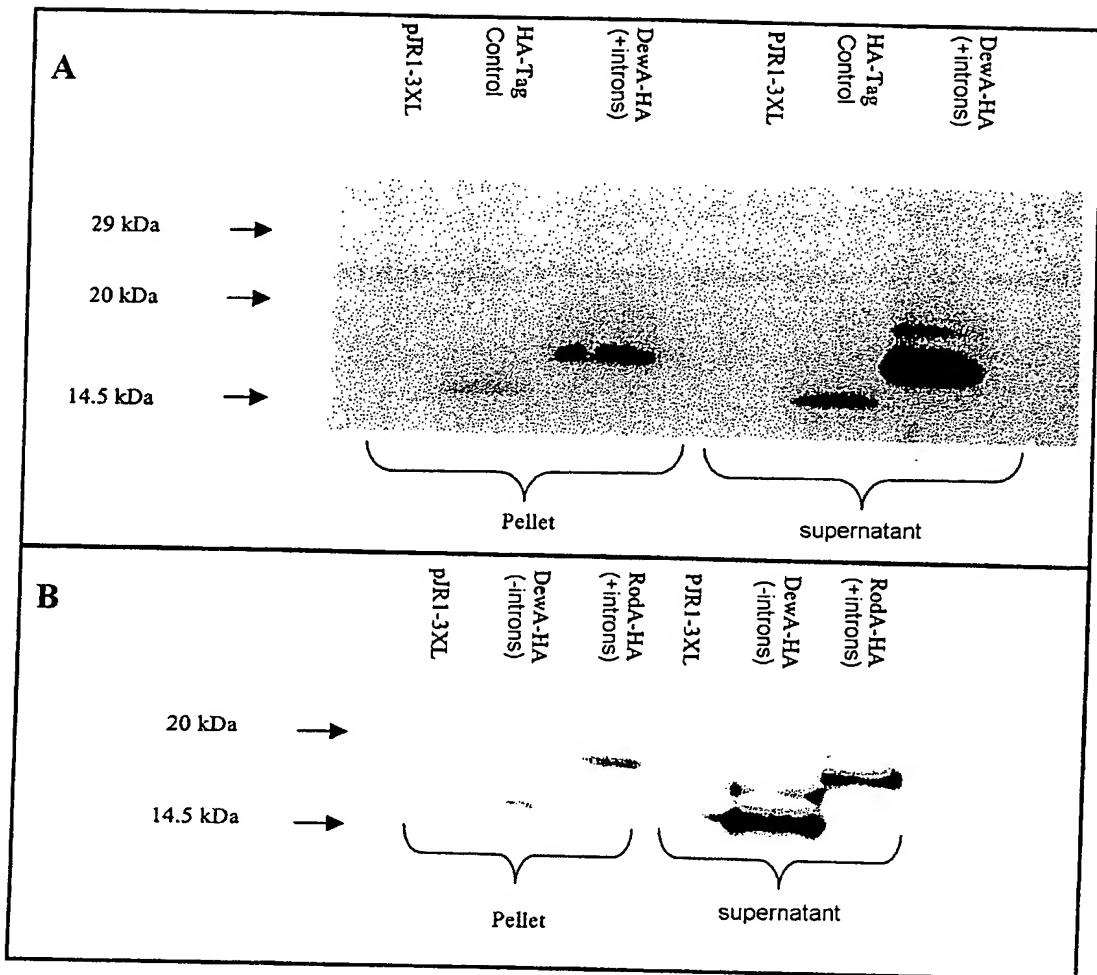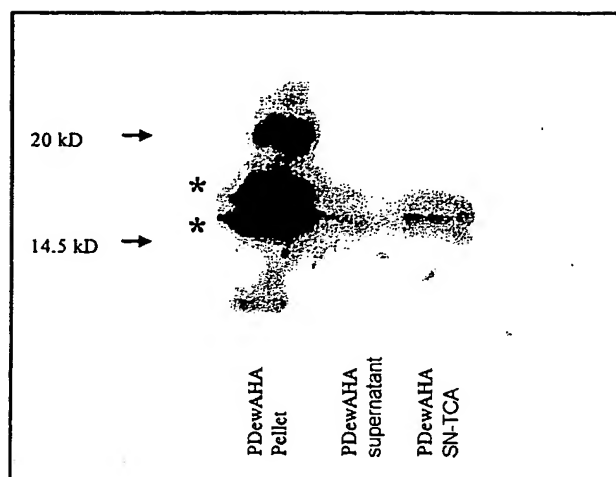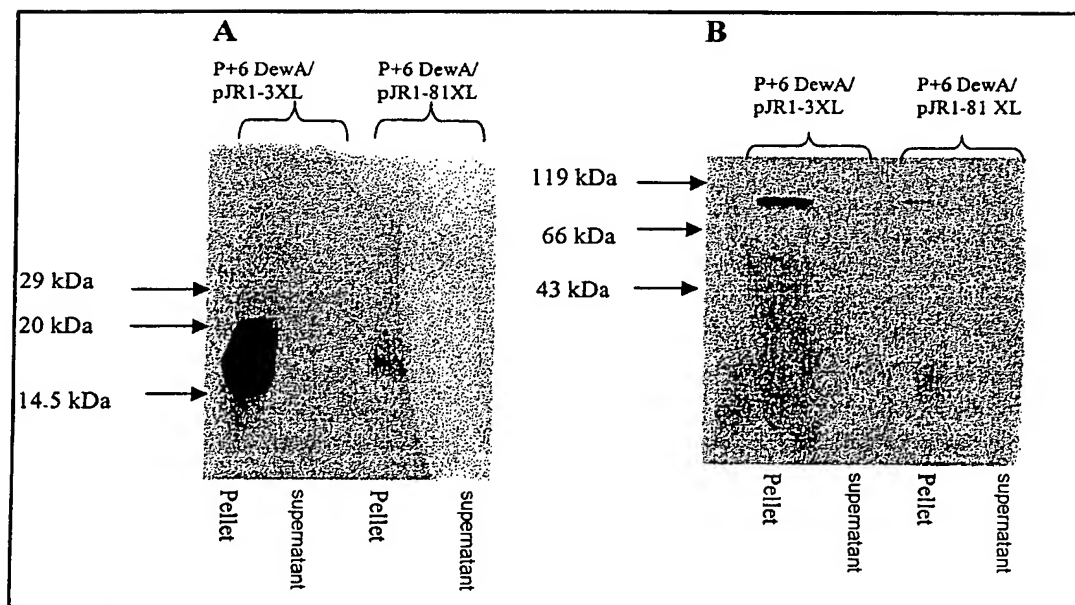CAGTGCGGCC GCTAATAG)

Fig.5

Fig.6

Fig.7

Fig.8

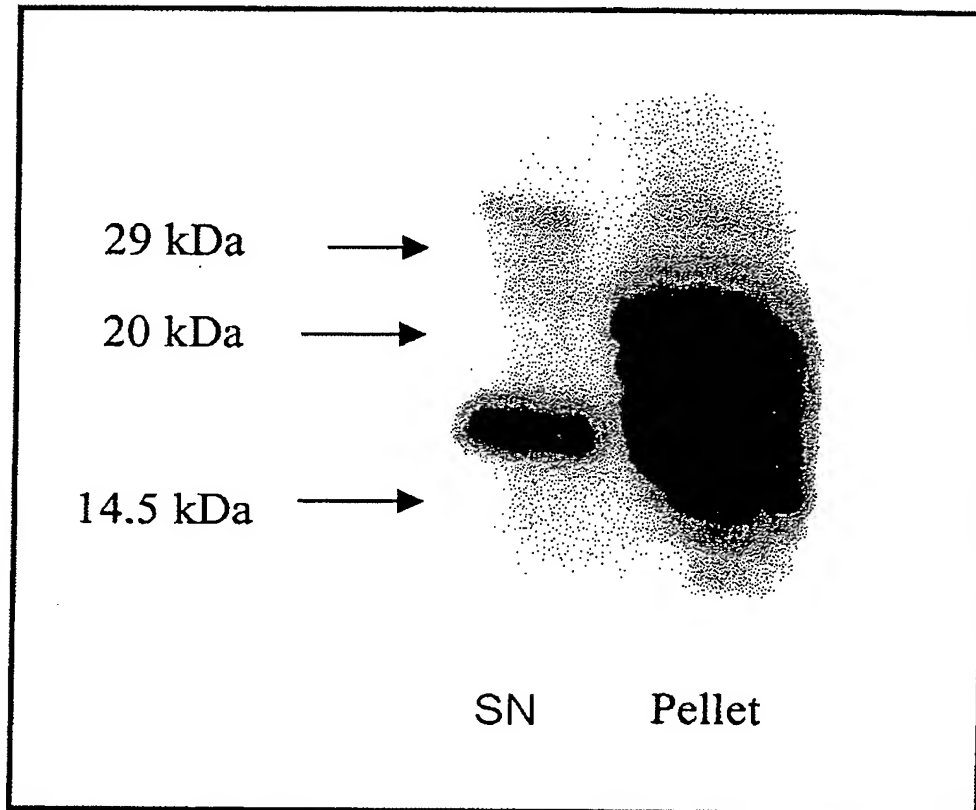29 kDa →

20 kDa →

14.5 kDa →

SN   Pellet

Fig. 9

A) *mfm1+* gene

Sequence of the mfm1-pre-protein

MDSMANSVSSSSVVNAGNKPAETLNKTVKNYTPKVPYMCVIA

Sequence of the mfm1+-gene

atggactcaa tggctaactc cgtttcttcc tcctctgtcg tcaacgctgg caacaagcct
gctgaaactc ttaacaagac cgttaagaat tataccccca aggttcctta catgtgtgtc attgcataa

mfm1 mature M-pheromone

YTPKVPYMC

DNA-Sequence of the mature mfm1 M-pheromone

tataccccca aggttcctta catgtgt

B) *mfm2+*-gene

Sequence of the mfm2-pre-protein

MDSIATNTHSSSIVNAYNNNPTDVVKTQNIKNYTPKVPYMCVIA

Sequence of the mfm2+-gene

atggactcca ttgcaactaa cactcattct tcatccattg tcaatgccta caacaacaat
cctaccgatg ttgtaaaaac tcaaaacatt aaaaattata ctccaaaggt tccttatatg
tgtgtaattg cttaa

mfm2 mature M-pheromone

YTPKVPYMC

DNA-Sequence of the mature mfm2 M-pheromone

tata ctccaaaggt tccttatatg tgt

C) *mfm3+*-gene

Sequence of the mfm3-pre-protein

MDSMANTVSSSVVNTGNKPSETLNKTVKNYTPKVPYMCVIA

Sequence of the mfm3+-gene

atggactcaa tggctaacac tgtttcttcc tccgtcgtta acactggcaa caagccttct
gaaactctta acaagactgt taagaattat acccccaagg ttccttacat gtgtgtcatt gcataa

mfm3 mature M-pheromone

YTPKVPYMC

DNA-Sequence of the mature mfm3 M-pheromone

tat accccaagg ttccttacat gtgt

Fig. 10

Genomic sequence of the *RodA* gene

```
ATGAAGTTCT  CCATTGCTGC  CGCTGTCGTT  GCTTTCGCCG  CCTCCGTCGC  GGCCCTCCCT  CCTGCCCATG
ATTCCCAGTT  CGCTGGCAAT  GGTGTTGGCA  ACAAGGGCAA  CAGCAACGTC  AAGTTCCCTG  TCCCCGAAAA
CGTGACCGTC  AAGCAGGCCT  CCGACAAGTG  CGGTGACCAG  GCCCAGCTCT  CTTGCTGCAA  CAAGGCCACG
TACGCCGGTG  ACACCACAAC  CGTTGATGAG  GGTCTTCTGT  CTGGTGCCCT  CAGCGGCCTC  ATCGGCGCCG
GGTCTGGTGC  CGAAGGTCTT  GGTCTCTTCG  ATCAGTGCTC  CAAGCTTGAT  GTTGCTGGTC  AGTTCTTCGA
AAATCACTTT  CGTGATGCCC  CAATGCTAAC  AATTACCAGT  CCTCATTGGC  ATCCAAGATC  TTGTCAACCA
GAAGTGCAAG  CAAAACATTG  CCTGCTGCCA  GAACTCCCCC  TCCAGCGCGG  TATGTTCCCT  TGTTTTACAG
CTTATTCACT  TAAACCGATT  AATCTAACAA  CGCTCACAGG  ATGGCAACCT  TATTGGTGTC  GGTCTCCCTT
GCGTTGCCCT  TGGCTCCATC  CTCTAA
```

DNA sequence of the open reading frame (ORF) of the *RodA* gene

```
ATGAAGTTCT  CCATTGCTGC  CGCTGTCGTT  GCTTTCGCCG  CCTCCGTCGC  GGCCCTCCCT  CCTGCCCATG
ATTCCCAGTT  CGCTGGCAAT  GGTGTTGGCA  ACAAGGGCAA  CAGCAACGTC  AAGTTCCCTG  TCCCCGAAAA
CGTGACCGTC  AAGCAGGCCT  CCGACAAGTG  CGGTGACCAG  GCCCAGCTCT  CTTGCTGCAA  CAAGGCCACG
TACGCCGGTG  ACACCACAAC  CGTTGATGAG  GGTCTTCTGT  CTGGTGCCCT  CAGCGGCCTC  ATCGGCGCCG
GGTCTGGTGC  CGAAGGTCTT  GGTCTCTTCG  ATCAGTGCTC  CAAGCTTGAT  GTTGCTGTCC  TCATTGGCAT
CCAAGATCTT  GTCAACCAGA  AGTGCAAGCA  AAACATTGCC  TGCTGCCAGA  ACTCCCCCTC  CAGCGCGGAT
GGCAACCTTA  TTGGTGTCGG  TCTCCCTTGC  GTTGCCCTTG  GCTCCATCCT  CTAA
```

Sequence of the RodA protein

```
MKFSIAAAVV  AFAASVAALP  PAHDSQFAGN  GVGNKGNSNV  KFPVPENVTV  KQASDKCGDQ  AQLSCCNKAT
YAGDTTTVDE  GLLSGALSGL  IGAGSGAEGL  GLFDQCSKLD  VAVLIGIQDL  VNQKCKQNIA  CCQNSPSSAD
GNLIGVGLPC  VALGSIL
```